

COINDUCTIVE PROOF PRINCIPLES FOR STOCHASTIC PROCESSES*

DEXTER KOZEN

Department of Computer Science, Cornell University, Ithaca, New York 14853-7501, USA
e-mail address: kozen@cs.cornell.edu

ABSTRACT. We give an explicit coinduction principle for recursively-defined stochastic processes. The principle applies to any closed property, not just equality, and works even when solutions are not unique. The rule encapsulates low-level analytic arguments, allowing reasoning about such processes at a higher algebraic level. We illustrate the use of the rule in deriving properties of a simple coin-flip process.

1. INTRODUCTION

Coinduction has been shown to be a useful tool in functional programming. Streams, automata, concurrent and stochastic processes, and recursive types have been successfully analyzed using coinductive methods; see [1, 7, 3, 11, 5] and references therein.

Most approaches emphasize the relationship between coinduction and bisimulation. In Rutten’s treatment [11] (see also [5, 1]), the coinduction principle states that under certain conditions, two bisimilar processes must be equal. For example, to prove the equality of infinite streams $\sigma = \text{merge}(\text{split}(\sigma))$, where **merge** and **split** satisfy the familiar coinductive definitions

$$\begin{aligned}\text{merge}(a :: \sigma, \tau) &= a :: \text{merge}(\tau, \sigma) \\ \#1(\text{split}(a :: b :: \rho)) &= a :: \#1(\text{split}(\rho)) \\ \#2(\text{split}(a :: b :: \rho)) &= b :: \#2(\text{split}(\rho)),\end{aligned}$$

it suffices to show that the two streams are bisimilar. An alternative view is that certain systems of recursive equations over a certain algebraic structure have unique solutions. Desharnais et al. [3, 7] study bisimulation in a probabilistic context. They are primarily interested in the approximation of one process with another. Again, they focus on bisimulation, but do not formulate an explicit coinduction rule.

In this paper we introduce a generalization of the coinduction principle that applies to other properties besides equations and to situations in which solutions are not unique. We illustrate its use with an extended example that demonstrates how the rule encapsulates low-level analytic arguments involving convergent sequences in its proof of soundness, thereby allowing reasoning about such processes at a higher algebraic level.

1998 ACM Subject Classification: F.4.1, F.3.1, I.1.3, I.2.3.

Key words and phrases: coinduction, coalgebra, logic in computer science, probabilistic logic, fractal.

* A preliminary version of this paper appeared as [9].

2. AN EXAMPLE

Consider the following procedure for simulating a coin of arbitrary real bias q , $0 \leq q \leq 1$, with a coin of arbitrary real bias p , $0 < p \leq 1/2$. We assume unit-time exact arithmetic on real numbers.

```

1  boolean QFLIP( $q$ ) {
2      if ( $q > p$ ) {
3          if (PFLIP()) return true;
4          else return QFLIP( $((q - p)/(1 - p))$ );
5      } else {
6          if (PFLIP()) return QFLIP( $q/p$ );
7          else return false;
8      }
9  }
```

Intuitively, if $q > p$ and the bias- p coin flip returns heads (**true**), which occurs with probability p , then we halt and output heads; this gives a fraction p/q of the desired probability q of heads of the simulated bias- q coin. If the bias- p coin returns tails, which occurs with probability $1 - p$, we rescale the problem appropriately and call QFLIP tail-recursively. Similarly, if $q \leq p$ and the bias- p coin returns tails, then we halt and output tails; and if not, we rescale appropriately and call QFLIP tail-recursively.

On any input $0 \leq q \leq 1$, the probability of halting is 1, since the procedure halts with probability at least p in each iteration. The probability that QFLIP halts and returns heads on input q exists and satisfies the recurrence

$$H(q) = \begin{cases} p \cdot H(\frac{q}{p}), & \text{if } q \leq p, \\ p + (1 - p) \cdot H(\frac{q-p}{1-p}), & \text{if } q > p. \end{cases} \quad (2.1)$$

Now $H^*(q) = q$ is a solution to this recurrence, as can be seen by direct substitution. There are uncountably many other solutions as well, but these are all unbounded (see Section 4). Since H^* is the unique bounded solution, it must give the probability of heads.

We can do the same for the expected running time. Let us measure the expected number of calls to PFLIP on input q . The expectation exists and is uniformly bounded on the unit interval by $1/p$, the expected running time of a Bernoulli (coin-flip) process with success probability p . From the program, we obtain the recurrence

$$\begin{aligned} E_0(q) &= \begin{cases} (1 - p) \cdot 1 + p \cdot (1 + E_0(\frac{q}{p})), & \text{if } q \leq p, \\ p \cdot 1 + (1 - p) \cdot (1 + E_0(\frac{q-p}{1-p})), & \text{if } q > p \end{cases} \\ &= \begin{cases} 1 + p \cdot E_0(\frac{q}{p}), & \text{if } q \leq p, \\ 1 + (1 - p) \cdot E_0(\frac{q-p}{1-p}), & \text{if } q > p. \end{cases} \end{aligned}$$

The unique bounded solution to this recurrence is

$$E_0^*(q) = \frac{q}{p} + \frac{1 - q}{1 - p}. \quad (2.2)$$

That it is a solution can be ascertained by direct substitution; uniqueness requires a further argument, which we will give later. As before, there are uncountably many unbounded solutions, but since E_0^* is the unique bounded solution, it must give the expected running time for any q .

The situation gets more interesting when we observe that slight modifications of the algorithm lead to noncontinuous fractal solutions with no simple characterizations like (2.2). The fractal behavior of stochastic processes has been previously observed in [6].

Currently, when $q > p$, we halt and output “heads” when PFLIP gives heads, which occurs with probability p . But note that we can save some time when $q \geq 1 - p$. In that case, we can halt and report heads if PFLIP gives tails, which occurs with the larger probability $1 - p$. This allows us to take off a larger fraction of the remaining “heads” weight of the bias- q coin. If PFLIP gives tails, we must still rescale, but the rescaling function is different. The new code is in lines 2–4.

```

1  boolean QFLIP( $q$ ) {
2      if ( $q \geq 1 - p$ ) {
3          if (PFLIP()) return QFLIP( $((q - (1 - p))/p)$ );
4          else return true;
5      } else if ( $q > p$ ) {
6          if (PFLIP()) return true;
7          else return QFLIP( $((q - p)/(1 - p))$ );
8      } else {
9          if (PFLIP()) return QFLIP( $q/p$ );
10         else return false;
11     }
12 }
```

The recurrence for the expected running time is

$$E_1(q) = 1 + r(q)E_1(f_1(q)), \quad (2.3)$$

where

$$f_1(q) = \begin{cases} \frac{q}{p}, & \text{if } q \leq p \\ \frac{q-p}{1-p}, & \text{if } p < q < 1-p \\ \frac{q-(1-p)}{p}, & \text{if } q \geq 1-p \end{cases} \quad (2.4)$$

$$r(q) = \begin{cases} 1-p, & \text{if } p < q < 1-p \\ p, & \text{otherwise.} \end{cases} \quad (2.5)$$

Again, there is a unique bounded solution

$$E_1^*(q) = \sum_{n=0}^{\infty} \prod_{j=0}^{n-1} r(f_1^j(q)),$$

but there is no longer a nice algebraic characterization like (2.2). The solution for $p = 1/4$ is the noncontinuous fractal shown in Fig. 1, shown compared to the straight line E_0^* running from $4/3$ to 4 . The large discontinuity at $q = 1 - p = 3/4$ is due to the modification of the algorithm for $q \geq 1 - p$, and this discontinuity is propagated everywhere by the recurrence.

Fig. 1 and intuition dictate that $E_1^* \leq E_0^*$, but how do we prove this? Not by induction, because there is no basis. One might briefly imagine that it is because the second process halts no later than the first on any predetermined sequence of coin flips, but there are trivial counterexamples. An analytic argument involving convergence of sequences seems inevitable.

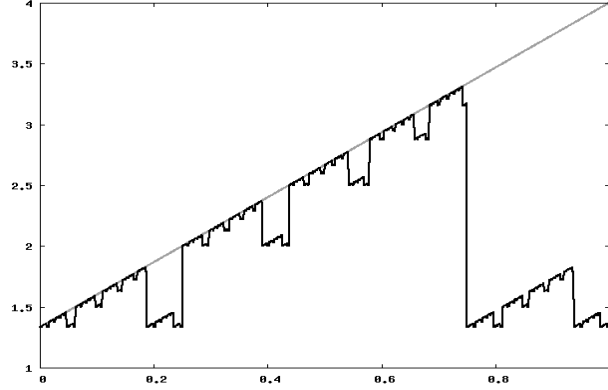


Figure 1: Fractal solution of (2.3)

However, there is a simpler alternative. It will follow from our coinductive proof principle that to conclude $E_1^* \leq E_0^*$, it suffices to show that $\tau(E_1)(q) \leq E_0^*(q)$ whenever $E_1(f_1(q)) \leq E_0^*(f_1(q))$, where τ is a suitably defined operator representing the unwinding of the recurrence (2.3) once. This property is easily checked algebraically, and no analysis is necessary.

We can modify the algorithm further to try to achieve more savings. If $1/2 < q < 1 - p$, it would seem to our advantage to remove p from the tail probability of q rather than from the head probability. The intuition behind this heuristic is that when q is in one of the regions $[0, p]$ or $[1 - p, 1]$, we can halt in the next step with the higher probability $1 - p$. If $q > 1/2$, then the proposed new action will cause q to move to the right toward the closer good region $[1 - p, 1]$ instead of to the left, thereby getting to a good region faster. The new code is in lines 5–7.

```

1  boolean QFLIP(q) {
2    if ( $q \geq 1 - p$ ) {
3      if (PFLIP()) return QFLIP( $(q - (1 - p))/p$ );
4      else return true;
5    } else if ( $q > 1/2$ ) {
6      if (PFLIP()) return false;
7      else return QFLIP( $q/(1 - p)$ );
8    } else if ( $q > p$ ) {
9      if (PFLIP()) return true;
10     else return QFLIP( $(q - p)/(1 - p)$ );
11   } else {
12     if (PFLIP()) return QFLIP( $q/p$ );
13     else return false;
14   }
15 }
```

The recurrence is

$$E_2(q) = 1 + r(q)E_2(f_2(q)) \quad (2.6)$$

with

$$f_2(q) = \begin{cases} \frac{q}{p}, & \text{if } q \leq p, \\ \frac{q-p}{1-p}, & \text{if } p < q \leq 1/2, \\ \frac{q}{1-p}, & \text{if } 1/2 < q < 1-p, \\ \frac{q-(1-p)}{p}, & \text{if } q \geq 1-p, \end{cases}$$

and $r(q)$ as given in (2.5). The symmetric fractal solution

$$E_2^*(q) = \sum_{n=0}^{\infty} \prod_{j=0}^{n-1} r(f_2^j(q))$$

is shown in Fig. 2.

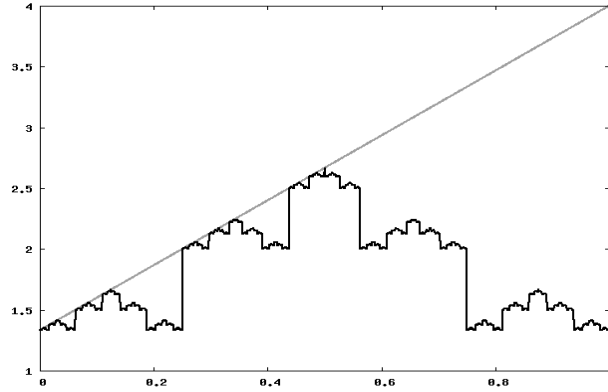


Figure 2: Fractal solution of (2.6)

Intuition seems to say that this solution should be at least as good as E_1^* , but it turns out that this is not always the case. By unwinding the recurrences a few steps and using the lower bound

$$E_2^*(q) \geq \sum_{n=0}^{\infty} p^n = \frac{1}{1-p},$$

it can be shown that for $p = 1/4$,

$$\begin{aligned} E_1^*(11/20) &= 5/2 = 2.5 \\ E_2^*(11/20) &\geq 323/128 \approx 2.5234375 \dots \end{aligned}$$

Moreover, this inversion holds on an open interval containing $11/20$ and countably many other open intervals.

One might ask whether there is a slight modification of E_2^* that is everywhere better than E_1^* . The answer is yes: take the breakpoint not at $1/2$, but at

$$c = \max((1-p)^2, 1-(1-p)^2),$$

provided $p \leq (1-p)^2$. For $p = 1/4$, this gives $c = 9/16$. Now the recurrence is

$$E_3(q) = 1 + r(q)E_3(f_3(q)) \tag{2.7}$$

with

$$f_3(q) = \begin{cases} \frac{q}{p}, & \text{if } q \leq p, \\ \frac{q-p}{1-p}, & \text{if } p < q \leq c, \\ \frac{q}{1-p}, & \text{if } c < q < 1-p, \\ \frac{q-(1-p)}{p}, & \text{if } q \geq 1-p. \end{cases}$$

(Since $p \leq (1-p)^2$ implies $p \leq (3 - \sqrt{5})/2 \approx 0.382$, this modification will not work for all p .)

Now we wish to show that $E_3^* \leq E_1^*$ on the whole unit interval. Note that we are comparing two nowhere-differentiable functions¹; we have no nice algebraic description of them save as solutions of the recurrences $E_i(q) = 1 + r(q)E_i(f_i(q))$. However, we can prove the desired inequality purely algebraically using the coinductive principle below, without recourse to analysis. We outline a proof below, after we have stated the principle and proved its validity.

3. A COINDUCTION PRINCIPLE

In this section we state and prove a coinduction principle that will allow us to derive properties of stochastic processes. The version we will use is most conveniently formulated in terms of bounded linear operators on a Banach space (complete normed linear space), but is closely related to a coinduction principle that holds in arbitrary complete metric spaces. We treat the metric version first.

Let (V, d) be a complete metric space. A function $\tau : V \rightarrow V$ is *contractive* if there exists a $c < 1$ such that for all $u, v \in V$, $d(\tau(u), \tau(v)) \leq c \cdot d(u, v)$. The value c is called the *constant of contraction*. A continuous function τ is said to be *eventually contractive* if τ^n is contractive for some $n \geq 1$. Contractive maps are uniformly continuous, and by the Banach fixpoint theorem, any such map has a unique fixpoint in V .

The fixpoint of a contractive map τ can be constructed explicitly as the limit of a Cauchy sequence $u, \tau(u), \tau^2(u), \dots$ starting at any point $u \in V$. The sequence is Cauchy; one can show by elementary arguments that

$$d(\tau^{n+m}(u), \tau^n(u)) \leq c^n(1 - c^m)(1 - c)^{-1} \cdot d(\tau(u), u).$$

Since V is complete, the sequence has a limit u^* , which by continuity must be a fixpoint of τ . Moreover, u^* is unique: if $\tau(u) = u$ and $\tau(v) = v$, then

$$d(u, v) = d(\tau(u), \tau(v)) \leq c \cdot d(u, v) \Rightarrow d(u, v) = 0,$$

therefore $u = v$.

Eventually contractive maps also have unique fixpoints. If τ^n is contractive, let u^* be the unique fixpoint of τ^n . Then $\tau(u^*)$ is also a fixpoint of τ^n . But then $d(u^*, \tau(u^*)) = d(\tau^n(u^*), \tau^{n+1}(u^*)) \leq c \cdot d(u^*, \tau(u^*))$, therefore u^* is also a fixpoint of τ .

In this framework, the coinduction rule takes the following simple form. If φ is a closed nonempty subset of a complete metric space V , and if τ is an eventually contractive map

¹Hermite and Poincaré eschewed such functions, calling them a “dreadful plague”. Poincaré wrote: “Yesterday, if a new function was invented, it was to serve some purpose; today, they are invented only to debunk the arguments of our predecessors, and they will never have any other use.”

on V that preserves φ , then the unique fixpoint u^* of τ is in φ . Expressed as a proof rule, this says for φ a closed property,

$$\frac{\exists u \varphi(u) \quad \forall u \varphi(u) \Rightarrow \varphi(\tau(u))}{\varphi(u^*)} \quad (3.1)$$

This is quite easily proved. Since φ is nonempty, it contains a point u . Since φ is preserved by τ , all elements of the sequence $\tau^n(u)$ are contained in φ . Finally, since φ is closed, the fixpoint u^* is contained in φ , since it is the limit of a Cauchy subsequence.

For our purposes, the coinduction principle is most conveniently expressed in the following form. This form makes clear how the principle allows analytic arguments to be replaced by simpler algebraic ones. See [4] for the necessary background.

Let B be a Banach space (complete normed linear space) over \mathbb{C} and let R be a bounded linear operator on B (*bounded* is synonymous with *continuous* for linear operators on B). The *spectrum* of R , denoted $\sigma(R)$, is the set of complex numbers λ such that $\lambda I - R$ is not invertible. The *spectral radius* of R is

$$\sup_{\lambda \in \sigma(R)} |\lambda| = \inf_n \sqrt[n]{\|R^n\|}, \quad (3.2)$$

where

$$\|R\| = \sup_{\|x\|=1} \|R(x)\|.$$

Suppose that $I - R$ is invertible; that is, $1 \notin \sigma(R)$. Let $a \in B$. Then there is a unique solution e^* of the equation $e = a + Re$, namely $e^* = (I - R)^{-1}a$.

Theorem 3.1. *Consider the affine operator $\tau(e) = a + Re$, where R is a bounded linear operator with spectral radius strictly less than 1. Let $\varphi \subseteq B$ be a closed nonempty region preserved by τ . Then $e^* \in \varphi$.*

Proof. By (3.2), if the spectral radius of R is less than 1, then R is eventually contractive; that is, there exists n such that $\|R^n\| < 1$. Then τ is also eventually contractive, since

$$\begin{aligned} \|\tau^n(e) - \tau^n(e')\| &= \left\| \sum_{i=0}^{n-1} R^i(a) + R^n(e) - \sum_{i=0}^{n-1} R^i(a) - R^n(e') \right\| \\ &= \|R^n(e - e')\| \\ &\leq \|R^n\| \cdot \|e - e'\|. \end{aligned}$$

It follows from (3.1) that the unique fixpoint of τ^n is contained in φ . But this fixpoint must be e^* , since e^* is a fixpoint of τ . \square

Restated as a proof rule, Theorem 3.1 takes the following form:

Theorem 3.2. *Let τ be as in Theorem 3.1. Let φ be a closed property. The following rule is valid:*

$$\frac{\exists e \varphi(e) \quad \forall e \varphi(e) \Rightarrow \varphi(\tau(e))}{\varphi(e^*)}. \quad (3.3)$$

More generally, for any $n \geq 1$,

$$\frac{\exists e \varphi(e) \quad \forall e \varphi(e) \Rightarrow \varphi(\tau^n(e))}{\varphi(e^*)}. \quad (3.4)$$

Proof. The rule (3.3) is just a restatement of Theorem 3.1. The rule (3.4) follows by applying (3.3) to the closed property $\psi(e) = \bigvee_{i=0}^{n-1} \varphi(\tau^i(e))$. This is a closed property because τ is continuous on B . \square

For example, to show that $E_1^* \leq E_0^*$ using the rule (3.3), we take B to be the space of bounded real-valued functions on the unit interval, $a = \lambda x.1$, $R : B \rightarrow B$ the bounded linear operator

$$R = \lambda E. \lambda q. r(q)E(f_1(q)) \quad (3.5)$$

with spectral radius $1 - p$, $\varphi(E)$ the closed property

$$\forall q \ E(q) \leq \frac{q}{p} + \frac{1-q}{1-p},$$

and

$$\tau(E) = \lambda q. (1 + r(q)E(f_1(q))) = \lambda q. (1 + RE(q)), \quad (3.6)$$

where f_1 and r are as given in (2.4) and (2.5). That the spectral radius of R is at most $1 - p$ follows immediately from (3.2), since

$$\|R\| = \sup_{\|E\|=1} \sup_q |RE(q)| = \sup_{\|E\|=1} \sup_q |r(q)E(f_1(q))| \leq 1 - p.$$

That it is exactly $1 - p$ requires a further argument, which we defer to Section 4.

Now the desired conclusion is

$$\forall q \ E_1^*(q) \leq \frac{q}{p} + \frac{1-q}{1-p}, \quad (3.7)$$

and the two premises we must establish are

$$\exists E \ \forall q \ E(q) \leq \frac{q}{p} + \frac{1-q}{1-p}, \quad (3.8)$$

$$\forall E \ \left(\forall q \ E(q) \leq \frac{q}{p} + \frac{1-q}{1-p} \Rightarrow \forall q \ \tau(E)(q) \leq \frac{q}{p} + \frac{1-q}{1-p} \right). \quad (3.9)$$

The premise (3.8) is trivial; for example, take $E = \lambda q.0$. For (3.9), let E be arbitrary. We wish to show that

$$\forall q \ E(q) \leq \frac{q}{p} + \frac{1-q}{1-p} \Rightarrow \forall q \ \tau(E)(q) \leq \frac{q}{p} + \frac{1-q}{1-p}. \quad (3.10)$$

Picking q arbitrarily on the right-hand side and then specializing the left-hand side at $f_1(q)$, it suffices to show

$$E(f_1(q)) \leq \frac{f_1(q)}{p} + \frac{1-f_1(q)}{1-p} \Rightarrow \tau(E)(q) \leq \frac{q}{p} + \frac{1-q}{1-p}. \quad (3.11)$$

Substituting the definition of τ , we need to show

$$E(f_1(q)) \leq \frac{f_1(q)}{p} + \frac{1-f_1(q)}{1-p} \Rightarrow 1 + r(q)E(f_1(q)) \leq \frac{q}{p} + \frac{1-q}{1-p}. \quad (3.12)$$

The proof breaks into three cases, depending on whether $q \leq p$, $p < q < 1 - p$, or $q \geq 1 - p$. In the first case, $f_1(q) = q/p$ and $r(q) = p$. Then (3.12) becomes

$$E\left(\frac{q}{p}\right) \leq \frac{q}{p^2} + \frac{1-\frac{q}{p}}{1-p} \Rightarrow 1 + pE\left(\frac{q}{p}\right) \leq \frac{q}{p} + \frac{1-q}{1-p}.$$

But

$$1 + pE\left(\frac{q}{p}\right) \leq 1 + p\left(\frac{q}{p^2} + \frac{1 - \frac{q}{p}}{1 - p}\right) = \frac{q}{p} + \frac{1 - q}{1 - p}.$$

The remaining two cases are equally straightforward. The last case, $q \geq 1 - p$, uses the fact that $p \leq 1/2$.

One can also prove closed properties of more than one function E . For example, as promised, we can show that $E_3^* \leq E_1^*$ whenever

$$\max((1 - p)^2, 1 - (1 - p)^2) \leq c \leq 1 - p.$$

For this application, B is the space of pairs (E, E') , where E and E' are bounded real-valued functions on the unit interval, $a = (\lambda x.1, \lambda x.1)$, and $R : B \rightarrow B$ is the bounded linear operator

$$R(E, E') = (\lambda q.r(q)E(f_3(q)), \lambda q.r(q)E'(f_1(q)))$$

with spectral radius $1 - p$. The closed property of interest is $E \leq E'$, but we need the stronger coinduction hypothesis

$$\varphi(E, E') = \forall q E(q) \leq E'(q) \tag{3.13}$$

$$\wedge E(q) \geq \frac{1}{1 - p} \tag{3.14}$$

$$\wedge p < q < 1 - p \Rightarrow E'(q) \geq 2 \tag{3.15}$$

$$\wedge E'(q) \leq \frac{q}{p} + \frac{1 - q}{1 - p} \tag{3.16}$$

$$\wedge 0 \leq q \leq p \Rightarrow E(q) = E(q + 1 - p). \tag{3.17}$$

Equivalent to (3.17) is the statement

$$1 - p \leq q \leq 1 \Rightarrow E(q) = E(q - (1 - p)). \tag{3.18}$$

There certainly exist (E, E') satisfying φ . We have also already argued that coinduction hypothesis (3.16) is preserved by τ . The argument for (3.14) is similar. For (3.17), if $0 \leq q \leq p$, then

$$1 - p \leq q + 1 - p \leq 1,$$

therefore

$$\begin{aligned} r(q) &= r(q + 1 - p) = p \\ f_3(q) &= \frac{q}{p} \\ f_3(q + 1 - p) &= \frac{(q + 1 - p) - (1 - p)}{p} = \frac{q}{p}. \end{aligned}$$

It follows that

$$1 + r(q)E(f_3(q)) = 1 + r(q + 1 - p)E(f_3(q + 1 - p)) = 1 + pE(q/p).$$

For (3.15), if $p < q < 1 - p$, then

$$\begin{aligned} r(q) &= 1 - p \\ E'(f_1(q)) &\geq \frac{1}{1 - p} \end{aligned}$$

by the coinduction hypotheses (3.13) and (3.14), thus

$$1 + r(q)E'(f_1(q)) \geq 1 + (1-p)\frac{1}{1-p} = 2.$$

Finally, for (3.13), we wish to show

$$1 + r(q)E(f_3(q)) \leq 1 + r(q)E'(f_1(q)),$$

or equivalently,

$$E(f_3(q)) \leq E'(f_1(q)). \quad (3.19)$$

Since f_1 and f_3 coincide except in the range $c < q < 1-p$, we need only show (3.19) for q in this range.

It follows from the assumptions in effect that

$$p < f_1(q) = \frac{q-p}{1-p} < 1-p < \frac{q}{1-p} = f_3(q),$$

thus

$$\begin{aligned} E(f_3(q)) &= E\left(\frac{q}{1-p} - (1-p)\right) && \text{by (3.17), in the form (3.18)} \\ &\leq \frac{\frac{q}{1-p} - (1-p)}{p} + \frac{1 - (\frac{q}{1-p} - (1-p))}{1-p} && \text{by (3.16)} \\ &= \left(\frac{q}{1-p} - 1\right)\frac{1-2p}{p(1-p)} + 2 \\ &\leq 2 && \text{since } p, q \leq 1-p \\ &\leq E'(f_1(q)) && \text{by (3.15).} \end{aligned}$$

We can conclude from the coinduction rule that $\varphi(E_3^*, E_1^*)$. Note that nowhere in this proof did we use any analytic arguments. All the necessary analysis is encapsulated in the proof of Theorem 3.1.

As a final application, we show how to use the coinductive proof rule (3.3) of Theorem 3.2 to argue that for $p < 1/2$, the function E_1^* is nowhere differentiable. We do this by showing that E_1^* has a dense set of discontinuities on the unit interval.

First we show that E_1^* has discontinuities at p and $1-p$. We know from clause (3.15) of the previous argument that for all q in the range $p < q < 1-p$,

$$E_1^*(q) \geq 2. \quad (3.20)$$

Also, by (3.7), we have that $E_1^*(q) \leq 1/p$ for all q . Then for $\varepsilon < p^2$, unwinding the defining recurrence (2.3) for E_1^* twice yields

$$E_1^*(1-p+\varepsilon) = 1+p+p^2E_1^*\left(\frac{\varepsilon}{p^2}\right) \leq 1+2p \quad (3.21)$$

$$E_1^*(p-\varepsilon) = 1+p+p^2E_1^*\left(1-\frac{\varepsilon}{p^2}\right) \leq 1+2p. \quad (3.22)$$

Since $1+2p < 2$, (3.20)–(3.22) imply that E_1^* has discontinuities at p and $1-p$.

Finally, we show that every nonempty open interval contains a discontinuity. Suppose for a contradiction that E_1^* is continuous on a nonempty open interval (a, b) . The interval (a, b) can contain neither p nor $1-p$, so the entire interval must be contained in one of the three regions $(0, p)$, $(p, 1-p)$, or $(1-p, 1)$.

Suppose it is contained in $(0, p)$. Then

$$E_1^*(q) = 1 + pE_1^*(q/p)$$

for $a < q < b$, thus

$$E_1^*(q/p) = (E_1^*(q) - 1)/p$$

for $a/p < q/p < b/p$, so E_1^* is also continuous on the interval $(a/p, b/p)$. But the length of this interval is $(b-a)/p$, thus we have produced a longer interval on which E_1^* is continuous.

A similar argument holds if (a, b) is contained in one of the intervals $(p, 1-p)$ or $(1-p, 1)$. In each of these three cases, we can produce an interval of continuity that is longer than (a, b) by a factor of at least $1/(1-p)$. This process can be repeated at most $\log(b-a)/\log(1-p)$ steps before the interval must contain one of the discontinuities p or $1-p$. This is a contradiction.

4. UNBOUNDED SOLUTIONS

That these coinductive proofs have no basis is reflected in the fact that there exist unbounded solutions in addition to the unique bounded solutions. All unbounded solutions are necessarily noncontinuous, because any continuous solution on a closed interval is bounded.

Theorem 3.1 does not mention these unbounded solutions, because they live outside the Banach space B . Nevertheless, it is possible to construct unbounded solutions to any of the above recurrences. All these recurrences are of the form

$$E(q) = a + r(q)E(f(q)). \quad (4.1)$$

Let G be the graph with vertices $q \in [0, 1]$ and edges $(q, f(q))$. Note that every vertex in G has outdegree 1. Let C be an undirected connected component of G . One can show easily that the following are equivalent:

- (i) C contains an undirected cycle;
- (ii) C contains a directed cycle;
- (iii) for some $q \in C$ and $k > 0$, $f^k(q) = q$.

Call C *rational* if these conditions hold of C , *irrational* otherwise. For example, for f_1 given in (2.4), the connected components of 0 and 1 are rational, since $f_1(0) = 0$ and $f_1(1) = 1$. There are other rational components besides these; for example, if $p = 1/4$, the component of $q = 11/20$ is rational, since $f_1^2(11/20) = f_1^4(11/20) = 1/5$.

Now any solution E of (4.1) must agree with the unique bounded solution E^* on the rational components: if $f^k(q) = q$, then unwinding the recurrence k times gives

$$E(q) = a \sum_{n=0}^{k-1} \prod_{i=0}^{n-1} r(f^i(q)) + \left(\prod_{i=0}^{k-1} r(f^i(q)) \right) E(q),$$

therefore

$$E(q) = \frac{a \sum_{n=0}^{k-1} \prod_{i=0}^{n-1} r(f^i(q))}{1 - \prod_{i=0}^{k-1} r(f^i(q))}.$$

But the values of E on an entire connected component are uniquely determined by its value on a single element of the component, since $E(q)$ uniquely determines $E(f(q))$ and vice-versa. Thus E and E^* must agree on the entire component.

We note in passing that this allows us to construct an E such that $RE = (1 - p)E$, where R is the linear operator of (3.5), thereby establishing that the spectral radius of R is $1 - p$. Take $E(1) = 1$, then inductively define $E(q) = r(q)E(f_1(q))/(1 - p)$ for all other q in the component of 1 and $E(q) = 0$ otherwise. Then $\|E\| = 1$, and

$$RE(q) = r(q)E(f_1(q)) = r(q)E(q) \cdot \frac{1 - p}{r(q)} = (1 - p)E(q).$$

For an irrational component, since there are no cycles, it is connected as a tree. We can freely assign an arbitrary value to an arbitrarily chosen element q of the component, then extend the function to the entire component uniquely and without conflict.

For $f \in \{f_1, f_2, f_3\}$ of the examples of Section 2, there always exists an irrational component. This follows from the fact that if $f^k(q) = q$, then q is a rational function of p ; that is, q is an element of the field $\mathbb{Q}(p)$. To see this, note that any $f^k(q)$ is of the form

$$\frac{q}{p^m(1 - p)^{k-m}} - r$$

for some $0 \leq m \leq k$ and $r \in \mathbb{Q}(p)$. This can be shown by induction on k . Solving $f^k(q) = q$ for q gives

$$q = \frac{rp^m(1 - p)^{k-m}}{1 - p^m(1 - p)^{k-m}} \in \mathbb{Q}(p).$$

Thus the component of any real $q \notin \mathbb{Q}(p)$ is an irrational component. There exist uncountably many such q , since $\mathbb{Q}(p)$ is countable. In fact, there are uncountably many irrational components, since each component is countable, and a countable union of countable sets is countable. Moreover, it can be shown that if q_1 and q_2 are in the same component, then $\mathbb{Q}(p, q_1) = \mathbb{Q}(p, q_2)$. This is because if q_1 and q_2 are in the same component, then $f^{k_1}(q_1) = f^{k_2}(q_2)$ for some $k_1, k_2 \in \mathbb{N}$, so

$$\frac{q_1}{p^{m_1}(1 - p)^{k_1-m_1}} - r_1 = \frac{q_2}{p^{m_2}(1 - p)^{k_2-m_2}} - r_2,$$

therefore $q_1 \in \mathbb{Q}(p, q_2)$ and $q_2 \in \mathbb{Q}(p, q_1)$.

We have thus characterized all possible solutions.

5. WHY IS THIS COINDUCTION?

The reader may be curious why we have called the rule (3.3) a coinduction rule, since it may seem different from the usual forms of coinduction found in the literature. The form of the rule and its use in applications certainly bears a resemblance to other versions in the literature, but to justify the terminology on formal grounds, we must exhibit a category of coalgebras and show that the rule (3.3) is equivalent to the assertion that a certain coalgebra is final in the category.

Say we have a contractive map τ on a metric space B and a nonempty closed subset $\varphi \subseteq B$ preserved by τ . Define $\tau(\varphi) = \{\tau(s) \mid s \in \varphi\}$. Consider the category C whose objects are the nonempty closed subsets of B and whose arrows are the reverse set inclusions; thus there is a unique arrow $\varphi_1 \rightarrow \varphi_2$ iff $\varphi_1 \supseteq \varphi_2$. The map $\bar{\tau}$ defined by $\bar{\tau}(\varphi) = \text{cl}(\tau(\varphi))$, where cl denotes closure in the metric topology, is an endofunctor on C , since $\bar{\tau}(\varphi)$ is a nonempty closed set, and $\varphi_1 \supseteq \varphi_2$ implies $\bar{\tau}(\varphi_1) \supseteq \bar{\tau}(\varphi_2)$. A $\bar{\tau}$ -coalgebra is then a nonempty closed set φ such that $\varphi \supseteq \bar{\tau}(\varphi)$; equivalently, such that $\varphi \supseteq \tau(\varphi)$. The final coalgebra is $\{e^*\}$, where

e^* is the unique fixpoint of τ . The coinduction rule (3.3) says that $\varphi \supseteq \tau(\varphi) \Rightarrow \varphi \supseteq \{e^*\}$, which is equivalent to the statement that $\{e^*\}$ is final in the category of $\bar{\tau}$ -coalgebras.

6. FUTURE WORK

There is great potential in the use of proof principles similar to those of Theorem 3.2 for simplifying arguments involving probabilistic programs, stochastic processes, and dynamical systems. Such rules encapsulate low-level analytic arguments, thereby allowing reasoning about such processes at a higher algebraic or logical level. A few such applications have been described in the theory of streams, Markov chains and Markov decision processes, and non-well-founded sets [10]. Other possible application areas are complex and functional analysis, the theory of linear operators, measure theory and integration, random walks, fractal analysis, functional programming, and probabilistic logic and semantics.

In particular, probabilistic programs can be modeled as measurable kernels $R(x, A)$, which can be interpreted as forward-moving measure transformers or backward-moving measurable function transformers [2, 8]. The expectation functions considered in this paper were uniformly bounded, but there are examples of probabilistic programs for which this is not true. It would be nice to find rules to handle these cases.

An intriguing open problem is whether the optimal strategy for the coin-flip process of Section 2 is decidable. Specifically, given rational p, q , $0 < p, q < 1$, and a flip of the bias- p coin, can we decide what action to take to minimize the expected running time? It is known that E_3^* is not necessarily optimal.

ACKNOWLEDGEMENTS

Thanks to Terese Damhøj Andersen, Lars Backstrom, Juris Hartmanis, Geoff Kozen, Prakash Panangaden, Jan Rutten, and the anonymous referees. This work was supported in part by ONR Grant N00014-01-1-0968 and by NSF grant CCF-0635028. The views and conclusions herein are those of the author and do not necessarily represent the official policies or endorsements of these organizations or the US government.

REFERENCES

- [1] Jon Barwise and Lawrence Moss. *Vicious Circles: On the Mathematics of Non-Wellfounded Phenomena*, volume 60 of *CSLI Lecture Notes*. Center for the Study of Language and Information (CSLI), Stanford University, 1996.
- [2] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labeled Markov processes. In *Proc. Int. Conf. Concurrency Theory (CONCUR'99)*. Springer-Verlag, 1999.
- [3] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. The metric analogue of weak bisimulation for probabilistic processes. In *Proc. Conf. Logic in Computer Science (LICS'02)*. IEEE, 2002.
- [4] Nelson Dunford and Jacob T. Schwartz. *Linear Operators: Part I: General Theory*. John Wiley, 1957.
- [5] Andrew D. Gordon. A tutorial on co-induction and functional programming. In *Proc. 1994 Glasgow Workshop on Functional Programming*. Springer Workshops in Computing, September 1994.
- [6] Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Stochastic processes as concurrent constraint programs. In *Proc. Conf. Principles of Programming Languages (POPL'99)*, pages 189–202, January 1999.
- [7] Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximate reasoning for real-time probabilistic processes. In *Proc. 1st Conf. Quantitative Evaluation of Systems (QEST'04)*. IEEE Press, 2004.
- [8] Dexter Kozen. Semantics of probabilistic programs. *J. Comput. Syst. Sci.*, 22:328–350, 1981.

- [9] Dexter Kozen. Coinductive proof principles for stochastic processes. In Rajeev Alur, editor, *Proc. 21st Symp. Logic in Computer Science (LICS'06)*, pages 359–366. IEEE, August 2006.
- [10] Dexter Kozen and Nicholas Ruozzi. Applications of metric coinduction. In T. Mossakowski et al., editor, *Proc. 2nd Conf. Algebra and Coalgebra in Computer Science (CALCO 2007)*, volume 4624 of *Lecture Notes in Computer Science*, pages 327–341. Springer, August 2007.
- [11] J.J.M.M. Rutten. Behavioural differential equations: a coinductive calculus of streams, automata, and power series. *Theoretical Computer Science*, 308:1–53, 2003.